Answer to reviewers - Second round
'Proper account of auto-correlations improves decoding performances of state-space (semi) Markov models.

Bez et al.

January 24, 2025

Recommander : Nathalie Peyrard

Decision
(minor) revision
The authors have answered in a very satisfying way almost all the comments of the first round of revision. Reviewer 2 asks for more clarifications on one of her initial comments. I follow this request and I describe below the reason.
Answers are provided in green. For tractability issues, the new modifications in the manuscript are made in red.

Review
Following the second revision of reviewer 2, I recommend to clarify in the protocol presentation and in the discussion, which comparisons are made and why. It was not obvious for me and reviewer 2 to recover this information, so I think a less attentive reader could miss it. In particular, I understand, but I may be wrong, that based on figure 9, bottom right, you compare the accuracy obtained with the HSMM-AR1 model on real data and the accuracy obtained with the mixture model of data simulated from a HSMM-AR1 model with parameters learned on the real data. I don't understand the reason of this comparison since both the data and the models are different.
Good point. For real data application, the comparison of accuracies between a HSMM-AR1model and a mixture model is not meant to be made but some elements of the manuscript (one sentence in the discussion and figure 9) tend to indicate it is done. This brought confusion as mentioned by the recommender and by reviewer 2.
We thus removed real case outputs from figure 9 and put them in an updated figure 13 so that we are now homogeneous in terms of what is where: models' description and accuracies of the simulation-estimation in figure 9; real data description and accuracies of states estimates in figure 13. We also reformulated the lines 474-477

Regarding the HSMM framework, there exists several versions of HSMM, from the more general to that described in Barbu and Limnios 2008 and to a simpler one referred to as ED-HMM or ED-HSMM. In the manuscript the authors describe the HSMM of Barbu and Limnios 2008. However the definition of the kernel is not as in this book. The probability of sojourn duration should be conditionally to the current state and the next state (see Barbu and Limnios p 3, or 45-46). However, in practice usually the user works with a simpler model where it depends only on the current state, this is the ED-HMM. I think in the experiments the model used in the ED-HMM, but this should be clarified.
Good point. This comment goes with one of reviewer 1. Indeed Barbu and Limnios in page 46 mention the particular case of a dependence on the current state alone. We clarified this in the new version by providing the correct general expression for HSMM (the final conditioning of the probability for the time duration was missing as pointed out by reviewer 1), and the particular one corresponding to our experiments. However, we did not mentioned the concept of Explicit Duration HMM (ED-HMM) to avoid adding a new concept.

Reviewer 1: Anonymous

I am very satisfied by the modifications made by the authors in response to my comments. I also like the new title for the article. I only have a few minor comments on the new parts of the manuscript, listed below, which should be very quick to address.

L.119: "and where" Did you mean while ?
Done
TWO LINES BELOW L.132: "on the jumps between the state process" Did you mean something like "on the process driving jumps between states" ? Overall the syntax of this sentence is a bit clumsy, to the point of obscuring the meaning.
Given that the difference between HMM and HSMM translate into the use of two different words i.e. a Markov process for HMM and a Markov chain for HSMM, we are a bit reluctant to follow the reformulation suggested here. To address the clumsy drawback of the sentence, we added two sentences to underline the types of memory that are addressed in HMM and in HSMM. We hope this will address the point made here.
P.7, FIRST EQUATION: Above the equation, you write that the sojourn time in the present state depends (only) on the present state, so shouldn't P (eTn+1 = k) be P (eTn+1 = k— eSn = s) ?
Good point that is connected to the second remark made by the recommender. We now i) provide the general expression of an HSMM and ii) its simplified version where the dependance is reduced to the current state.
"DISCRIMINATION BETWEEN PDFS", L.6: "the state is no longer considered (...) and the state estimator"
Done
P.11, L.7: "insure" ⇒ "ensures" ?
Done.

Reviewer 2: Sandra Plancade, 07 Jan 2025

My original comment
On real data, accuracy is less good with HSMM-AR1 than expected in a mixture model analysis in setting 2 (Figure 9). This observation is mentioned on l.429-430, but without explanation. The most surprising is that this occures for setting/vessel 2 only, while model mispecification is greater with setting/vessel 1. Do you have an explanation or hypotheses?

Response of the authors
Before answering, we felt the need to clear up a misunderstanding in the reviewer's comment. We believe the reviewer swapped the two boats because the comment is very specific for vessel 1 (settings 1) and not for vessel 2 (settings 2) . On this basis, i) For real data, accuracy is less good with HSMM-AR1 than expected in a mixture model analysis in 3 setting 1 (and not in settings 2 as indicated by the reviewer; blue is for setting 1 and red for setting 2). ii) lines 429-430 concern the model performance under simulations and not the real data. We rather believe that reviewer 2 drew our attention on lines 461-470 that concern her point.

First, we found that line 460 needed reformulation. The new formulation is now : For Vessel 1, the most robust model (HSMM - AR1) produces state decoding performances that under-perform the simulation experiments and the mixture model. However, for vessel 2, it produces performances between the simulation experiments and the mixture model. Second, as mentioned by the reviewer, the mispecification is greater with setting 1. Our explanation of the bad performances of HSMM-AR1 for settings 1 are thus provided in lines 468-470 (old numbering). It can be summarized in a confusion in estimated parameters related 1) to less distinguishable speed distributions between the two states and 2) to high variability of the ns parameter of the Negative Binomial distribution.

My response to the response
All apologies for the messing up in my comment about vessel 1/2 and the line numbers, and thank you very much for having put it back in order.

:)

Nevertheless, I am not fully convinced by your answer. Indeed, my question was about the underperformances of HSMM-AR1 with respect to the mixture model for vessel 1 (blue), in Figure 9, bottom-right.

See above the answer about mixture models to the recommender question: we modified figures 9 and 13.

However, your first argument of less distinguishable speed distributions between the two states explains the worse decoding performances on data from vessel 1 w.r.t. vessel 2, but I don't understand how it is related to the superiority of mixture model on HSMM-AR1 for vessel 1. I consider that this point should be clarified.

Good point. This was clumsy. This was polished (see the new formulation lines 477-479).

Your second argument mentioned in your answer above, namely the high variability of the ns parameter of the NB distribution that quantifies the deviation from the geometric distribution, seems different from the one presented in the MS (old numbering 468-470 / new numbering 479-481) related to the AR order. I probably have misunderstood something...

Good point. We fully agree. We mixed up several things in our answer that should have not been mixed. The second argument is the one of the manuscript: the speed data of vessel 1 get autocorrelation features that are the most distant from the model hypothesis.

Moreover, I wonder if the arguments suggested in the following lines (from new numbering 482) in the form of a recommendation, namely a mispecification in the correlation between observations that goes beyond AR order, are part of the explanation of the comparative performances of HSMM-AR1 and mixture model.

They are part of an explanation of the bad performances notably for vessel 1. Given that the reference to mixture model made above in the text has now been removed, we believe this remark vanishes also.

More generally, my questionning probably originate from a lack of understanding about the role of the comparison with the mixture model. Indeed, while the models HMM/HSMM-AR0/AR1 are partially nested, I don't understand precisely the positionning of the mixture model. If it is explained somewhere else in the MS, I apologize for having missed it but it would be relevant to refer to it when analysing the performances of the mixture model (around lines 471-...). If not, I recommend to clarify this point.

We agree that there was a small confusion in the manuscript. The motivation to refer to mixture model is explained in the section on the "Discrimination between speed PDF". It can be rephrased as following. If the discrimination between speed distributions is such that one can estimate the state with very good confidence from the knowledge of the sole speed values, then the situation is favorable. This discriminating power is measured by d_V which serves as a baseline for state decoding accuracies. According to this baseline, settings 2 are clearly more favorable than settings 1.

In the simulation-estimation experiments, the idea behind the comparison between H(S)MM and mixture model state decoding accuracies is the following. Complex models like HMM or HSMM incorporate temporal state dynamic that (are expected to) help improving state decoding. Improving with regards to what? Improving with regard to what can be done without honoring the existence of batches of states induced by the Markov properties of the model. This later situation is handled by mixture models where states are estimated by the sole knowledge of the speed values. Improvement is achieved when the data are consistent with the model structure as indicated by the simulation-estimation experiments. See lines 436-439 where this point is raised.